



41

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Hewick, Rodney M.
Wang, Jack H.
Wozney, John M.
Celeste, Anthony J.

(ii) TITLE OF INVENTION: Bone and Cartilage Inductive Proteins

(iii) NUMBER OF SEQUENCES: 15

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Legal Affairs, Genetics Institute, Inc.
(B) STREET: 87 CambridgePark Drive
(C) CITY: Cambridge
(D) STATE: MA
(E) COUNTRY: USA
(F) ZIP: 02140

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/319,831
(B) FILING DATE: 06-OCT-1994
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Kapinos, Ellen J.
(B) REGISTRATION NUMBER: 32,245
(C) REFERENCE/DOCKET NUMBER: GI 5182A-DIV

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 617-876-1170
(B) TELEFAX: 617-876-5851

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(F) TISSUE TYPE: Bone

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Arg His Glu Leu Tyr Val Ser Phe Gln Asp Leu Gly Trp Leu Asp Trp
1 5 10 15
Val Ile Ala Pro Gln Gly Tyr
20

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Bos taurus
(F) TISSUE TYPE: Bone

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Leu Ser Ala Thr Ser Val Leu Tyr Tyr Asp Ser Ser Asn Asn Val Ile
1 5 10 15
Leu Arg

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Bos taurus*

(F) TISSUE TYPE: Bone

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Ala Cys Cys Ala Pro Thr Lys

1

5

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Bos taurus*

(F) TISSUE TYPE: Bone

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Thr Asn Glu Leu Pro Pro Pro Asn Lys Leu Pro Gly Ile Phe Asp Asp

1

5

10

15

Val His Gly Ser His Gly Arg

20

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 80 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Bos taurus

(vii) IMMEDIATE SOURCE:
(B) CLONE: acc30

(viii) POSITION IN GENOME:
(C) UNITS: bp

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 25..57

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGATCCGCGT GCTGTGCTCC GACC AAG GTG AGC GCC ACC TCC GTG CTC TAC 51
Lys Leu Ser Ala Thr Ser Val Leu Tyr
1 5

TAC GAC AGCAGCAACA ATGTAATTCT AGA 80
Tyr Asp
10

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Lys Leu Ser Ala Thr Ser Val Leu Tyr Tyr Asp
1 5 10

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 199 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Bos taurus

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Bovine genomic
(B) CLONE: Lambda 9800-10

(viii) POSITION IN GENOME:

(C) UNITS: bp

(ix) FEATURE:

- (A) NAME/KEY: exon
(B) LOCATION: 30..199

(ix) FEATURE:

- (A) NAME/KEY: intron
(B) LOCATION: 1..29

(ix) FEATURE:

- (A) NAME/KEY: CDS
(B) LOCATION: 30..179

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TGCCCCGCTGC CCCCTCCCGC CCCCGCCAG GTG CAC CTG CTG AAG CCG CAC GCG 53
Val His Leu Leu Lys Pro His Ala
1 5

GTC CCC AAG GCG TGC TGC GCG CCC ACC AAG CTG AGC GCC ACT TCC GTG 101
Val Pro Lys Ala Cys Cys Ala Pro Thr Lys Leu Ser Ala Thr Ser Val
10 15 20

CTC TAC TAC GAC AGC AGC AAC AAC GTC ATC CTG CGC AAG CAC CGC AAC 149
Leu Tyr Tyr Asp Ser Ser Asn Asn Val Ile Leu Arg Lys His Arg Asn
25 30 35 40

ATG GTG GTC CGC GCC TGC GGC TGC CAC TGAGGCCCA ACTCCACCGG 196
Met Val Val Arg Ala Cys Gly Cys His
45 50

CAG 199

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 49 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Val His Leu Leu Lys Pro His Ala Val Pro Lys Ala Cys Cys Ala Pro
 1 5 10 15
 Thr Lys Leu Ser Ala Thr Ser Val Leu Tyr Tyr Asp Ser Ser Asn Asn
 20 25 30
 Val Ile Leu Arg Lys His Arg Asn Met Val Val Arg Ala Cys Gly Cys
 35 40 45
 His

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 172 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Bos taurus

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Bovine genomic
- (B) CLONE: Lambda 9800-10

(viii) POSITION IN GENOME:

- (C) UNITS: bp

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 51..161

(ix) FEATURE:

- (A) NAME/KEY: intron
- (B) LOCATION: 1..50

(ix) FEATURE:

- (A) NAME/KEY: intron
- (B) LOCATION: 162..172

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 51..161

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GGGGTGGGAG GGCACGTGGA TGGGACTCAC CTTCTCCCAC TACCCCCCAG GAC TGG 56

Asp Trp

1

GTC ATC GCC CCC CAA GGC TAC TCA GCC TAT TAC TGT GAA GGG GAG TGC 104

Val Ile Ala Pro Gln Gly Tyr Ser Ala Tyr Tyr Cys Glu Gly Glu Cys

5

10

15

TCC TTC CCG CTG GAC TCC TGC ATG AAC GCC ACC AAC CAC GCC ATC CTG 152

Ser Phe Pro Leu Asp Ser Cys Met Asn Ala Thr Asn His Ala Ile Leu

20

25

30

CAG TCC CTG GTCAGTACCT C 172

Gln Ser Leu

35

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 37 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Asp Trp Val Ile Ala Pro Gln Gly Tyr Ser Ala Tyr Tyr Cys Glu Gly

1

5

10

15

Glu Cys Ser Phe Pro Leu Asp Ser Cys Met Asn Ala Thr Asn His Ala

20

25

30

Ile Leu Gln Ser Leu

35

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 119 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Bos taurus*

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: Bovine genomic

(B) CLONE: Lambda 9800-10

(viii) POSITION IN GENOME:

(C) UNITS: bp

(ix) FEATURE:

(A) NAME/KEY: exon

(B) LOCATION: 20..99

(ix) FEATURE:

(A) NAME/KEY: intron

(B) LOCATION: 1..19

(ix) FEATURE:

(A) NAME/KEY: intron

(B) LOCATION: 100..119

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 22..99

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CCCTTGCGTG TCCCCGCAGA C GAC GTC CAC GGC TCC CAC GGC CGG CAG GTG 51

Asp Val His Gly Ser His Gly Arg Gln Val

1

5

10

TGC CGT CGG CAC GAG CTG TAC GTG AGC TTC CAG GAC CTG GGC TGG CTG 99

Cys Arg Arg His Glu Leu Tyr Val Ser Phe Gln Asp Leu Gly Trp Leu

15

20

25

GTGAGTTCCG ACTCTCCTTT

119

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Asp Val His Gly Ser His Gly Arg Gln Val Cys Arg Arg His Glu Leu
 1 5 10 15

Tyr Val Ser Phe Gln Asp Leu Gly Trp Leu
 20 25

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1003 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens
- (F) TISSUE TYPE: Human Heart

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Human heart cDNA library stratagene catalog
#936208
- (B) CLONE: hH38

(viii) POSITION IN GENOME:

- (C) UNITS: bp

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 8..850

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 427..843

(ix) FEATURE:

- (A) NAME/KEY: mRNA
- (B) LOCATION: 1..997

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GAATTCC GAG CCC CAT TGG AAG GAG TTC CGC TTT GAC CTG ACC CAG ATC
 Glu Pro His Trp Lys Glu Phe Arg Phe Asp Leu Thr Gln Ile
 -139 -135 -130

CCG GCT GGG GAG GCG GTC ACA GCT GCG GAG TTC CGG ATT TAC AAG GTG 97
Pro Ala Gly Glu Ala Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys Val
-125 -120 -115 -110

CCC AGC ATC CAC CTG CTC AAC AGG ACC CTC CAC GTC AGC ATG TTC CAG 145
Pro Ser Ile His Leu Leu Asn Arg Thr Leu His Val Ser Met Phe Gln
-105 -100 -95

GTG GTC CAG GAG CAG TCC AAC AGG GAG TCT GAC TTG TTC TTT TTG GAT 193
Val Val Gln Glu Gln Ser Asn Arg Glu Ser Asp Leu Phe Phe Leu Asp
-90 -85 -80

CTT CAG ACG CTC CGA GCT GGA GAC GAG GGC TGG CTG GTG CTG GAT GTC 241
Leu Gln Thr Leu Arg Ala Gly Asp Glu Gly Trp Leu Val Leu Asp Val
-75 -70 -65

ACA GCA GCC AGT GAC TGC TGG TTG CTG AAG CGT CAC AAG GAC CTG GGA 289
Thr Ala Ala Ser Asp Cys Trp Leu Leu Lys Arg His Lys Asp Leu Gly
-60 -55 -50

CTC CGC CTC TAT GTG GAG ACT GAG GAT GGG CAC AGC GTG GAT CCT GGC 337
Leu Arg Leu Tyr Val Glu Thr Glu Asp Gly His Ser Val Asp Pro Gly
-45 -40 -35 -30

CTG GCC GGC CTG CTG GGT CAA CGG GCC CCA CGC TCC CAA CAG CCT TTC 385
Leu Ala Gly Leu Leu Gly Gln Arg Ala Pro Arg Ser Gln Gln Pro Phe
-25 -20 -15

GTG GTC ACT TTC TTC AGG GCC AGT CCG AGT CCC ATC CGC ACC CCT CGG 433
Val Val Thr Phe Phe Arg Ala Ser Pro Ser Pro Ile Arg Thr Pro Arg
-10 -5 1

GCA GTG AGG CCA CTG AGG AGG AGG CAG CCG AAG AAA AGC AAC GAG CTG 481
Ala Val Arg Pro Leu Arg Arg Arg Gln Pro Lys Lys Ser Asn Glu Leu
5 10 15

CCG CAG GCC AAC CGA CTC CCA GGG ATC TTT GAT GAC GTC CAC GGC TCC 529
Pro Gln Ala Asn Arg Leu Pro Gly Ile Phe Asp Asp Val His Gly Ser
20 25 30 35

CAC GGC CGG CAG GTC TGC CGT CGG CAC GAG CTC TAC GTC AGC TTC CAG 577
His Gly Arg Gln Val Cys Arg Arg His Glu Leu Tyr Val Ser Phe Gln
40 45 50

GAC CTT GGC TGG CTG GAC TGG GTC ATC GCC CCC CAA GGC TAC TCA GCC 625
Asp Leu Gly Trp Leu Asp Trp Val Ile Ala Pro Gln Gly Tyr Ser Ala
55 60 65

TAT TAC TGT GAG GGG GAG TGC TCC TTC CCG CTG GAC TCC TGC ATG AAC 673
Tyr Tyr Cys Glu Gly Glu Cys Ser Phe Pro Leu Asp Ser Cys Met Asn
70 75 80

GCC ACC AAC CAC GCC ATC CTG CAG TCC CTG GTG CAC CTG ATG AAG CCA 721
 Ala Thr Asn His Ala Ile Leu Gln Ser Leu Val His Leu Met Lys Pro
 85 90 95

AAC GCA GTC CCC AAG GCG TGC TGT GCA CCC ACC AAG CTG AGC GCC ACC 769
 Asn Ala Val Pro Lys Ala Cys Cys Ala Pro Thr Lys Leu Ser Ala Thr
 100 105 110 115

TCT GTG CTC TAC TAT GAC AGC AGC AAC AAC GTC ATC CTG CGC AAG CAC 817
 Ser Val Leu Tyr Tyr Asp Ser Ser Asn Asn Val Ile Leu Arg Lys His
 120 125 130

CGC AAC ATG GTG GTC AAG GCC TGC GGC TGC CAC TGAGTCAGCC CGCCCAGCCC 870
 Arg Asn Met Val Val Lys Ala Cys Gly Cys His
 135 140

TACTGCAGCC ACCCTTCTCA TCTGGATCGG GCCCTGCAGA GGCAGAAAAC CCTTAAATGC 930

TGTCACAGCT CAAGCAGGAG TGTCAGGGGC CCTCACTCTC GGTGCCTACT TCCTGTCAGG 990

CTTCTGGGAA TTC 1003

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 281 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Glu Pro His Trp Lys Glu Phe Arg Phe Asp Leu Thr Gln Ile Pro Ala
 -139 -135 -130 -125

Gly Glu Ala Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys Val Pro Ser
 -120 -115 -110

Ile His Leu Leu Asn Arg Thr Leu His Val Ser Met Phe Gln Val Val
 -105 -100 -95

Gln Glu Gln Ser Asn Arg Glu Ser Asp Leu Phe Phe Leu Asp Leu Gln
 -90 -85 -80

Thr Leu Arg Ala Gly Asp Glu Gly Trp Leu Val Leu Asp Val Thr Ala
 -75 -70 -65 -60

Ala Ser Asp Cys Trp Leu Leu Lys Arg His Lys Asp Leu Gly Leu Arg
 -55 -50 -45

Leu Tyr Val Glu Thr Glu Asp Gly His Ser Val Asp Pro Gly Leu Ala
-40 -35 -30

Gly Leu Leu Gly Gln Arg Ala Pro Arg Ser Gln Gln Pro Phe Val Val
-25 -20 -15

Thr Phe Phe Arg Ala Ser Pro Ser Pro Ile Arg Thr Pro Arg Ala Val
-10 -5 1 5

Arg Pro Leu Arg Arg Arg Gln Pro Lys Lys Ser Asn Glu Leu Pro Gln
10 15 20

Ala Asn Arg Leu Pro Gly Ile Phe Asp Asp Val His Gly Ser His Gly
25 30 35

Arg Gln Val Cys Arg Arg His Glu Leu Tyr Val Ser Phe Gln Asp Leu
40 45 50

Gly Trp Leu Asp Trp Val Ile Ala Pro Gln Gly Tyr Ser Ala Tyr Tyr
55 60 65

Cys Glu Gly Glu Cys Ser Phe Pro Leu Asp Ser Cys Met Asn Ala Thr
70 75 80 85

Asn His Ala Ile Leu Gln Ser Leu Val His Leu Met Lys Pro Asn Ala
90 95 100

Val Pro Lys Ala Cys Cys Ala Pro Thr Lys Leu Ser Ala Thr Ser Val
105 110 115

Leu Tyr Tyr Asp Ser Ser Asn Asn Val Ile Leu Arg Lys His Arg Asn
120 125 130

Met Val Val Lys Ala Cys Gly Cys His
135 140

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 112 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Asp Val His Gly Ser His Gly Arg Gln Val Cys Arg Arg His Glu Leu
1 5 10 15

Tyr Val Ser Phe Gln Asp Leu Gly Trp Leu Leu Leu Val Ile Ala Pro
20 25 30

Gln Gly Tyr Ser Ala Tyr Tyr Cys Glu Gly Glu Cys Ser Phe Pro Leu
35 40 45

Asp Ser Cys Met Asn Ala Thr Asn His Ala Ile Leu Gln Ser Leu Val
50 55 60

His Leu Leu Lys Pro His Ala Val Pro Lys Ala Cys Cys Ala Pro Thr
65 70 75 80

Lys Leu Ser Ala Thr Ser Val Leu Tyr Tyr Asp Ser Ser Asn Asn Val
85 90 95

Ile Leu Arg Lys His Arg Asn Met Val Val Arg Ala Cys Gly Cys His
100 105 110